arXiv:1104.1242v1 [math.PR] 7 Apr 2011

Once more on comparison of tail index estimators

Vygantas Paulauskas[1,2] and Marijus Vaičiulis[2]

[1] Vilnius University, Department of Mathematics and Informatics,

[2] VU Institute of Mathematics and Informatics, Vilnius

January 19, 2013

Abstract

We consider heavy-tailed distributions and compare the well-known estimators of the tail index, based on extreme value theory with a comparatively recent estimator based on a different idea.

Short title: Comparison of tail index estimators

MSC 2000 subject classifications. Primary 62F12, secondary 62G32, 60F05 .

Key words and phrases. Estimation of tail index, Hill estimator, Pickands estimator

Corresponding author: Vygantas Paulauskas, Department of Mathematics and Informatics, Vilnius university, Naugarduko 24, Vilnius 03225, Lithuania, e-mail:vygantas.paulauskas@mif.vu.lt

1 Introduction

During last several decades it was demonstrated that in many fields of applied probability the so-called heavy-tailed distributions play an important role. One of the main problems, connected with heavy-tailed distributions is the estimation of the tail index - a parameter, which characterizes the heaviness of the tail of a distribution. The problem can be formulated as follows. Let us consider a sample $X_1, \dots, X_N$ of size N taken from a heavy-tailed distribution function (d.f.) F, that is, we assume that $X_1, \dots, X_N$ are independent identically distributed (i.i.d.) random variables with a d.f. F satisfying the following relation for large x:

$$1 - F(x) = x^{-\alpha} L(x). \tag{1}$$

Here $\alpha > 0$, $L(x) > 0$ for all $x > 0$ and$L$ is a slowly varying at infinity function:

$$\lim_{x \to \infty} \frac{L(tx)}{L(x)} = 1.$$

If we have only condition (1) without any additional information about the function L, it is difficult to get good properties, such as the asymptotic normality, of an estimator of the parameter α. Therefore the main stream of papers dealing with the tail index estimation uses the so-called second-order condition of regular variation . During last years even the third order condition on F was introduced (see, for example Fraga Alves et al(2006)). In our paper, like in Paulauskas (2003), we shall use the second order condition in the form of the relation (3) below.

Let $X_{N,1} \le X_{N,2} \le \cdots \le X_{N,N}$ denote the ordered statistics of $X_1, \dots, X_N$. Most of tail index estimators are based on ordered statistics and estimate the parameter $\gamma = 1/\alpha$. One of the most popular estimators to estimate the parameter $\gamma = 1/\alpha$ was proposed by Hill (1975):

$$\gamma^{(1)}_{N,k} = \frac{1}{k} \sum_{i=0}^{k-1} \log X_{N,N-i} - \log X_{N,N-k},$$

where k is some number satisfying $1 \le k \le N$. We also list some other

estimators based on ordered statistics:

$$\gamma_{N,k}^{(2)} = (\log 2)^{-1} \log \frac{X_{N,N-[k/4]} - X_{N,N-[k/2]}}{X_{N,N-[k/2]} - X_{N,N-k}},$$
$$\gamma_{N,k}^{(3)} = \gamma_{N,k}^{(1)} + 1 - \tfrac{1}{2}\big(1 - (\gamma_{N,k}^{(1)})^2/M_N\big)^{-1},$$
$$\gamma_{N,k}^{(4)} = \frac{M_N}{2\gamma_{N,k}^{(1)}},$$

where

$$M_N = \frac{1}{k}\sum_{i=0}^{k-1}(\log X_{N,N-i} - \log X_{N,N-k})^2.$$

The estimator $\gamma_{N,k}^{(2)}$ was proposed by Pickands (1975) , $\gamma_{N,k}^{(3)}$ – in Dekkers et al (1989) and $\gamma_{N,k}^{(4)}$ – by C. G. de Vries (see e.g. de Haan and Peng (1998)). There are many papers devoted to the modifications of the estimators $\gamma_{N,k}^{(i)}$, $i = 1, 2, 3, 4$. Let us mention several of them (in chronological order): Weissman (1978), Smith (1987), Resnick and Starica (1997), Geluk and Peng (2000), Fraga Alves (2001), Gomes and Martins (2002), Fraga Alves et al (2003), Li et al (2008), Gomes et al (2008).

All estimators $\gamma_{N,k}^{(i)}$, $i = 1, 2, 3, 4$ contain one additional parameter k, which has clear intuitive meaning in the case of all above written estimators $\gamma_{N,k}^{(i)}$: how many the largest values from the ordered statistics must be taken in order to have good properties (consistency, asymptotic normality, etc) of the estimator. We presented these four estimators because in de Haan and Peng (1998) all these estimators are compared and it is shown that none of these estimators dominates the others. It turned out that for different values of the parameters γ and ρ (the parameters characterizing the so-called second-order asymptotic behavior of F, which will be introduced below) different estimators have the smallest asymptotic mean-squared error.

In Davydov et al (2000) (see also Davydov and Paulauskas (1999)) there was proposed a new estimator, based on a different idea, which came when considering rather abstract objects – random compact convex stable sets. Although originally in Davydov et al (2000) an estimator was constructed to estimate the index of a multivariate stable distribution (even there was the restriction $0 < \alpha < 1$ for this index) and the main tool in the proof was the relation between exponential distribution and ordered statistics, in Paulauskas (2003) it was noted that the same construction of the

estimator can be based on a different idea and that this idea can be employed in the context of a general tail index estimation. We shall return to this point after introducing this new estimator. The construction of the estimator is as follows.

We divide the sample into n groups $V_1, \dots, V_n$, each group containing m random variables, that is, we assume that $N = n \cdot m$ and $V_i = \{X_{(i-1)m+1}, \dots, X_{im+1}\}$. (In practice at first m is chosen and then $n = [N/m]$ is taken, where $[x]$ stands for the integer part of a number $x > 0$.) Let

$$M_{ni}^{(1)} = \max\{X_j \colon X_j \in V_i\}$$

and let $M_{ni}^{(2)}$ denote the second largest element in the same group V_i. Let us denote

$$\kappa_{ni} = \frac{M_{ni}^{(2)}}{M_{ni}^{(1)}}, \qquad S_n = \sum_{i=0}^{n} \kappa_{ni}, \qquad Z_n = n^{-1} S_n. \tag{2}$$

From now instead of (1) we require stronger condition. Let us assume that we have a sample $X_1, \dots, X_N$ from distribution F, which satisfies the second-order asymptotic relation (as $x \to \infty$)

$$1 - F(x) = C_1 x^{-\alpha} + C_2 x^{-\beta} + \mathrm{o}(x^{-\beta}), \tag{3}$$

with some parameters $0 < \alpha < \beta \leq \infty$. It seems that Hall (see Hall (1982)) was the first who considered this condition in the context of tail index estimation and, under this condition, proved asymptotic normality of Hill estimator. Note that the case $\beta = \infty$ corresponds to Pareto distribution, $\beta = 2\alpha$ – to stable distribution with exponent $0 < \alpha < 2$. In de Haan and Peng (1998) (and in many other papers dealing with tail index estimation as well) more general second-order asymptotic relation is used in a different form with parameters $\gamma = 1/\alpha$ and $\rho = \alpha - \beta$ and in a more general context of the extreme-value index, when the parameter γ can take negative values, too. Namely, let U denotes the right continuous inverse of the function $1/(1-F)$. Suppose that there exits a function $A(t)$, which ultimately has constant sign and tends to zero as $t \to \infty$, then the relation

$$\lim_{x \to \infty} \frac{\frac{U(tx)}{U(t)} - x^{\gamma}}{A(t)} = x^{\gamma} \frac{x^{\rho} - 1}{\rho}, \quad \rho \leq 0, \tag{4}$$

serves for the definition of the second order regularly varying tail $1 - F$.

As it was mentioned above, the estimator Z_n from (2) (in a different context of a sample from multivariate stable distribution) was based on the following relation (see LePage et al (1981))

$$\big(M_{ni}^{(1)}, M_{ni}^{(2)}\big) m^{-1/\alpha} \xrightarrow{\mathrm{D}}_{N\to\infty} (\Gamma_1^{-1/\alpha}, \Gamma_2^{-1/\alpha}),$$

where $\Gamma_i = \sum_{j=1}^{i} \lambda_j$ and $\lambda_j, j \geq 1$, are i.i.d. standard exponential random variables and $\xrightarrow{\mathrm{D}}$ denotes convergence in distribution, and the fact that

$$\mathbf{E}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{1/\alpha} = \frac{\alpha}{1+\alpha}.$$

Here may be it is worth to mention that most estimation of tail index procedures, starting from Hill(1975) and Pickands (1975) papers, are based on the relation between order statistics and exponential distributions (the Renyi representation theorem): if $X_1, \dots, X_n$ is a sample from a continuous strictly increasing d.f. F, $F(0) = 0$, and $X^{(1)} \geq \cdots \geq X^{(n)}$ is the order statistics, then

$$X^{(i)} = F^{-1}\left(\exp\left\{-\sum_{j=1}^{i} \lambda_j (n-j+1)^{-1}\right\}\right), \quad i = 1, 2, \dots, n$$

and

$$\lambda_j = (n-j+1)\big(\ln F(X^{(j-1)}) - \ln F(X^{(j)})\big), \quad j = 1, 2, \dots, n$$

In Paulauskas (2003) it was noted that estimator from (2) can be based on a different idea. If we take two independent random variables X and Y with the same Pareto distribution

$$F(x) = 1 - C_1 x^{-\alpha}, \quad x \geq C_1^{1/\alpha},$$

and denote

$$W = \frac{\min(X,Y)}{\max(X,Y)}, \tag{5}$$

then it is not difficult to verify that, denoting $p = \alpha/(1+\alpha)$, we have $\mathbf{E}\,W = p$ (since W is invariant under scale transformation, we can take $C_1 = 1$ and in the sequel we shall refer to that case as a standard Pareto distribution). Therefore in the case of the Pareto distribution quantity Z_n, as an estimator

for the parameter p (we shall denote it by $\hat{p}$ and as in Qi (2010) we shall call it as DPR estimator), is nothing but the sample mean for a bounded random variable, moreover, in this case the best choice is to take m=2. If the underlying distribution F is not Pareto, but satisfies (3), then it is natural to expect that for large m $\mathbf{E}\,\hat{p} = \mathbf{E}\,\kappa_{n1}$ will be close to p. In Paulauskas (2003) the following estimate , which was the main ingredient in the proof of the asymptotic normality of the estimator $\hat{p}$, was given (see Lemma in Paulauskas (2003))

$$|\gamma_m| \le C_0 m^{-\zeta}, \tag{6}$$

where $\gamma_m = \mathbf{E}\,\hat{p} - p$, $\zeta = (\beta - \alpha)/\alpha$, and C_0 is a constant depending on C_1, C_2, α, and β. The DPR estimator is a sum of i.i.d. random variables, therefore $\mathbf{E}\,\hat{p} = \mathbf{E}\,\kappa_{n1}$ and γ_m gives the bias of our estimator. Based on (6) the following result was proved in Paulauskas (2003).

Theorem 1. *Let us suppose that F satisfies* (2) *with $0 < \alpha < \beta < \infty$. If we choose*

$$n = \varepsilon_N N^{2\zeta/(1+2\zeta)}, \qquad m = \varepsilon_N^{-1} N^{1/(1+2\zeta)},$$

where $\varepsilon_N \to 0$, as $N \to \infty$, then

$$\sqrt{n}(\hat{p} - p) \xrightarrow{\mathrm{D}}_{N\to\infty} N(0, \sigma^2), \tag{7}$$

where $\sigma^2 = \lim_{n\to\infty} \sigma_n^2 = \alpha((\alpha+1)^2(\alpha+2))^{-1}$.

Now we can give the exact asymptotic behavior of γ_m, this allows us to choose m in an optimal way and to compare DPR estimator with these four estimators, listed above, in the same manner as it was done in de Haan and Peng (1998) (therefore in the title of the paper there are words “once more”). Our main result can be formulated as follows. (We write $a_n \sim b_n$ if $\lim_{n\to\infty} a_n b_n^{-1} = 1$.)

Theorem 2. *Let us suppose that F satisfies* (2) *with $0 < \alpha < \beta < \infty$ and $C_1 > 0$. Then we have*

$$\gamma_m \sim \chi m^{-\zeta}, \quad (m \to \infty), \tag{8}$$

where

$$\chi = \chi(C_1, C_2, \alpha, \beta) = \frac{C_2 \beta \zeta \Gamma(\zeta+1)}{C_1^{\zeta+1}(\alpha+1)(\beta+1)}.$$

For sufficiently large N (ensuring that $m_{opt} \geq 2$) taking

$$m_{opt} = \left(\frac{2\zeta\chi^2}{\sigma^2}\right)^{1/(1+2\zeta)} N^{1/(1+2\zeta)} \tag{9}$$

we get that MSE (mean square error) is minimal

$$\mathbf{E}\ (\hat{p} - p)^2 \sim (1 + 2\zeta) \left(\frac{\chi^2\sigma^{4\zeta}}{(2\zeta)^{2\zeta}N^{2\zeta}}\right)^{1/(1+2\zeta)}. \tag{10}$$

Under this choice of m we have asymptotic normality

$$\sqrt{n}(\hat{p} - p) \xrightarrow{\mathrm{D}}_{N\to\infty} N(\mu, \sigma^2), \tag{11}$$

where σ^2 is the same as in Theorem 1 and $\mu = \sigma(2\zeta)^{-1/2}\ \mathrm{sgn}(\chi)$.

Remark 3. The estimator $\hat{p}$ (as all other introduced above tail index estimators) is invariant with respect to scale transformation, while condition (3) is not: if a random variable X_1 satisfy this condition, a random variable AX_1, where $A > 0$ with distribution function F_A satisfies the relation

$$1 - F_A(x) = C_1A^\alpha x^{-\alpha} + C_2A^\beta x^{-\beta} + \mathrm{o}(x^{-\beta}),$$

therefore the constants in the second order relation are not invariant. But it is easy to see that the ratio C_1^β/C_2^α is invariant and all quantities in relations (8), (9), and (10) depend exactly only on this ratio.

As it was mentioned, this result allows to compare the estimator $\hat{p}$ with four estimators mentioned above, and this is done in Section 3. Although according to the chosen criteria Hill estimator $\gamma_{N,k}^{(1)}$ and estimator $\gamma_{N,k}^{(4)}$ asymptotically perform better than the estimator $\hat{p}$, relation between other two estimators and $\hat{p}$ is the same as in paper de Haan and Peng (1998): for some values of parameters α, β estimator $\hat{p}$ performs better than $\gamma_{N,k}^{(2)}$ and $\gamma_{N,k}^{(3)}$. But here it is worth to mention the simple structure of the DPR estimator and the fact that it is well suited for recursive calculations (for example, when we have the so-called tick-by-tick financial data and we need tail index estimation in real time), see the monograph Markovich (2007). There are situations (such examples are mentioned in Oi(2010)) when data can be divided naturally into blocks but only few of largest observations within blocks

are available. In such situations estimator $\hat{p}$ can be applied while all other mentioned estimators are not applicable. Also the estimator $\hat{p}$ is well adapted for detecting a change in tail index, see a paper GADEIKIS and PAULAUSKAS (2005), where estimator $\hat{p}$ was used to analyze financial crisis in Asian markets in 1997-98 and the results were compared with analogous analysis using Hill estimator in QUINTOS et al(2001). And the main factor why we think that DPR estimator deserves the attention of statisticians is the possibility of several promising modifications of the estimator. One such modification is to introduce an additional parameter $r > 0$ and to consider the estimator

$$\hat{p}_r = n^{-1} S_{n,r}, \quad S_{n,r} = \sum_{i=1}^{n} \kappa_{ni}^r.$$

Again, using standard Pareto distribution, it is easy to calculate that $\hat{p}_r$ estimates the quantity $\alpha(r+\alpha)^{-1}$. As a matter of fact, when preparing the paper PAULAUSKAS (2003) the first named author had considered the estimator $\hat{p}_2$ (which to some extent resembles the quantity M_N, see definition of estimators $\gamma_{N,k}^{(3)}$ and $\gamma_{N,k}^{(4)}$), but realized that there is no gain in changing the first moment by the second one. Now it turns out that it is worth to take $0 < r < 1$, and we are able to prove asymptotic normality (under appropriate assumptions on m) for a fixed r. Also we can show that between two estimators with fixed parameters $0 < r' < r'' \le 1$, the smaller asymptotic MSE has $\hat{p}_{r'}$. Unfortunately, at present we do not know how to choose optimally r, which in general can be dependent on α, β, and even N.

Let us consider general construction of modifications of the DPR estimator. Take some function $f : [0,1] \to [0,\infty]$ such that $\mathbf{E}\, f(W)$ exists where W is from (5), then this expectation will be some function of α and, of course, on function f. Let us denote this function by $h_f(\alpha)$, that is $h_f(\alpha) = \mathbf{E}\, f(W)$. If h_f is a one-to-one map from $[a,b]$ to $[c,d]$ with $[a,b]$ and $[c,d]$ being subsets of $[0,\infty]$, then estimating the quantity $h_f(\alpha)$ and taking the inverse function we get an estimator for α (with the restriction $a \le \alpha \le b$ if $0 < a < b < \infty$). Therefore it is natural to consider statistic of the form

$$\frac{1}{n}\sum_{i=1}^{n} f(\kappa_{ni}), \tag{12}$$

obtaining large class of modifications of the estimator $\hat{p}$, developed in PAULAUSKAS (2003). The estimator $\hat{p}$ is obtained taking $f_1(x) = x$, then

$h_{f_1}(\alpha) = \alpha/(1+\alpha)$. The above mentioned modification $\hat{p}_r$ is obtained by taking $f_r(x) = x^r,\ r > 0$ and $h_{f_r}(\alpha) = h_r(\alpha) = \alpha/(r+\alpha)$. Estimators of the type (12) we shall call generalized DPR estimators, in short GDPR.

In a recent paper Qi (2010) one more estimator is introduced, which can be considered as connecting ideas of DPR and Hill estimators At first the procedure is the same as for DPR estimator - division of the sample in n groups with m elements in each group. But then instead of taking two largest elements in each group Qi takes Hill estimator in each group, namely, taking $s+1$ $(1 \le s \le m-1)$ largest values in each group, then averaging them over groups and obtaining the following estimator of the parameter $\gamma = \alpha^{-1}$:

$$\gamma_N(s) = \frac{1}{ns} \sum_{i=1}^{n} \sum_{j=1}^{s} \big(\log M_{ni}^{(j)} - \log M_{ni}^{(s+1)} \big), \tag{13}$$

where $M_{ni}^{(1)} \ge \ldots M_{ni}^{(m)}$ is ordered statistic from V_i. With $s = 1$ estimator (13) becomes of the form (12) with $f_\ell(x) = -\log x$. It is not difficult to calculate that for this function f we get $h_\ell(\alpha) = \alpha^{-1}$.

Having possibility to choose several functions in construction of GDPR estimators, natural question is what properties of these functions ensures better results in estimating α. Comparing two functions $f_1(x) = x$ and $f_\ell(x) = -\log x$ we see that corresponding functions $h_1(\alpha) = \alpha/(1+\alpha)$ and $h_\ell(\alpha) = \alpha^{-1}$ have quite different ranges: the first one has a small range - interval $(0,1)$, while the second one as a range have all half line $(0,\infty)$. Moreover, this fact results in different behavior of derivatives of inverse functions

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}p} h_1^{-1}(p) &= \frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{p}{1-p}\right) = \frac{1}{(1-p)^2} = (\alpha+1)^2, \\ \frac{\mathrm{d}}{\mathrm{d}\gamma} h_\ell^{-1}(\gamma) &= \frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{1}{\gamma}\right) = -\frac{1}{\gamma^2} - = \alpha^2. \end{aligned}$$

For small values of α (this corresponds to small values of p and large values of γ) the derivative of the first function is almost one, while for the second function it tends to zero as γ^{-2}. This means that even big changes in the value of estimated quantity γ results only in small changes of estimated value of α. For large values of α (as $p \to 1$ or $\gamma \to 0$) behavior of both derivatives is the same, but, evidently, large values of α are not so interesting in the problem of tail index estimation. These considerations explain why Qi estimator (13) with $s = 1$ (or, in other words, GDPR estimator with the

function f_ℓ) performs better than DPR estimator (with the function f_1) and also suggest one more modification of DPR estimators. Namely, if we take the same function f_r, but with negative parameter r (to stress this we shall write $f_{-r},\ r>0$), there will appear restriction $\alpha > r$, but now the range of the function $h_{-r}(\alpha) = \alpha(\alpha-r)^{-1}$ is infinite interval $(1,\infty)$ and the behavior of the derivative of inverse function is very similar to that of h_ℓ^{-1}. We are able to show that in the case of function f_{-r} it is possible to find optimal choice of r and GDPR estimator with this optimal r is comparable with estimator (13) with $s=1$ in a sense that for some values of $\alpha,\ \beta$ one estimator has smaller asymptotic MSE, for other values - dominates another one. Investigation of all these modifications were carried while the first version of the paper was in the process of refereeing and the results with proofs are collected in a forthcoming paper Paulauskas and Vaičiulis (2010).

One more remark concerning Theorem 2 is appropriate here. In Qi (2010) it is mentioned that using the same method of the proof of asymptotic normality for estimator (13)(that is, using the relation between ordered statistics and exponential distributions) it is possible to prove (11). Our proof of (11) essentially differs since it does not use exponential distributions and the main tool in the proof is formula (14). It is worth to mention that one can get the results for estimator (13) with $s=1$ by using this formula. This is demonstrated in the above mentioned forthcoming paper.

The paper contains two more sections. In Section 2 we prove Theorem 2 and in Section 3 there are results on comparison of estimators.

2 Proof of Theorem 2

Proof of (8). Relation (8) gives the exact asymptotic of the bias $\mathbf{E}\,\hat{p}-p$. Generally, the exact asymptotic of the bias of a tail index estimator is rather difficult problem. The advantage of our estimator $\hat{p}$ is a relative simplicity of the proof of (8). We do not use asymptotic for the inverse function for $1-F(x)$ as in Paulauskas (2003), but rather simple form of the expectation

$$\mathbf{E}\,\hat{p} \;=\; 1-m\int_0^\infty F^{m-1}(x_2)\left\{\int_{x_2}^\infty \frac{\mathrm{d}F(x_1)}{x_1}\right\}\mathrm{d}x_2, \tag{14}$$

which will be proved below. We truncate the outer integral at the level

$$a_m = \kappa m^{1/\alpha}\,(\ln m)^{-1/\alpha}\,, \tag{15}$$

where $0 < \kappa < (C_1/\zeta)^{1/\alpha}$ and denote

$$\begin{aligned} K_{m,1} &= m \int_0^{a_m} F^{m-1}(x_2) \left\{ \int_{x_2}^{\infty} \frac{\mathrm{d}F(x_1)}{x_1} \right\} \mathrm{d}x_2, \\ K_{m,2} &= m \int_{a_m}^{\infty} F^{m-1}(x_2) \left\{ \int_{x_2}^{\infty} \frac{\mathrm{d}F(x_1)}{x_1} \right\} \mathrm{d}x_2, \end{aligned}$$

Now, (8) follows immediately from the following two relations

$$K_{m,1} = o\left(m^{-\zeta}\right), \tag{16}$$

$$1 - p - K_{m,2} \sim \chi m^{-\zeta}, \quad m \to \infty. \tag{17}$$

To prove (16), split $K_{m,1}$ into two parts: $K_{m,1} = K'_{m,1} + K''_{m,1}$, where

$$K'_{m,1} = m \int_0^{a_m} F^{m-1}(x_2) \left\{ \int_{x_2}^{a_m} \frac{\mathrm{d}F(x_1)}{x_1} \right\} \mathrm{d}x_2.$$

By the change of integration order to get

$$\begin{aligned} K'_{m,1} &= m \int_0^{a_m} \frac{\mathrm{d}F(x_1)}{x_1} \left\{ \int_0^{x_1} F^{m-1}(x_2) \right\} \mathrm{d}x_2 \\ &\leq m \int_0^{a_m} F^{m-1}(x_1)\, \mathrm{d}F(x_1) = F^m(a_m). \end{aligned}$$

An assumption (3) and a simple inequality $\ln(1-x) \leq -x$, $0 \leq x < 1$ yield

$$F^m(a_m) \leq C \left(1 - C_1 a_m^{-\alpha}\right)^m \leq C \mathrm{e}^{m \ln(1 - C_1 a_m^{-\alpha})} \leq C m^{-C_1 \kappa^{-\alpha}}$$

for sufficiently large m, hence, taking into account (15), we get $K'_{m,1} = o\left(m^{-\zeta}\right)$. Relations $\int_0^{a_m} F^{m-1}(x)\, \mathrm{d}x = O\left(a_m F^{m-1}(a_m)\right)$ and $\int_{a_m}^{\infty} x^{-1}\, \mathrm{d}F(x) = O\left(a_m^{-\alpha-1}\right)$ prove

$$K''_{m,1} = O\left(m a_m^{-\alpha} F^{m-1}(a_m)\right) = o\left(m^{-\zeta}\right),$$

and we have (16).

Now let us prove (17). Integrating by parts the inner integral we get

$$K_{m,2} = m \int_{a_m}^{\infty} F^{m-1}(x) \left\{ \frac{C_1 \alpha}{\alpha+1} x^{-\alpha-1} + \frac{C_2 \beta}{\beta+1} x^{-\beta-1} + o\left(x^{-\beta-1}\right) \right\} \mathrm{d}x.$$

Denote $\tilde{F}(x) = 1 - C_1 x^{-\alpha} - C_2 x^{-\beta}$, then one can write

$$K_{m,2} = K'_{m,2} + K''_{m,2} + R_m,$$

where

$$\begin{aligned} K'_{m,2} &= \frac{m}{\alpha+1} \int_{a_m}^{\infty} \tilde{F}^{m-1}(x) \, \mathrm{d}\tilde{F}(x), \\ K''_{m,2} &= C_2 \beta m \left(\frac{1}{\beta+1} - \frac{1}{\alpha+1} \right) \int_{a_m}^{\infty} x^{-\beta-1} \tilde{F}^{m-1}(x) \, \mathrm{d}x, \end{aligned} \tag{18}$$

and

$$R_m = R_{m,1} + R_{m,2} \tag{19}$$

with

$$\begin{aligned} R_{m,1} &= m \int_{a_m}^{\infty} \left(F^{m-1}(x) - \tilde{F}^{m-1}(x) \right) \left\{ \frac{C_1 \alpha}{\alpha+1} x^{-\alpha-1} + \frac{C_2 \beta}{\beta+1} x^{-\beta-1} \right\} \mathrm{d}x, \\ R_{m,2} &= m \int_{a_m}^{\infty} F^{m-1}(x) o\left(x^{-\beta-1} \right) \, \mathrm{d}x. \end{aligned}$$

Integration gives $1 - p - K'_{m,2} = O\left(\tilde{F}^m(a_m) \right) = o\left(m^{-\zeta} \right)$. We claim that

$$K''_{m,2} \sim -\chi m^{-\zeta}, \quad (m \to \infty). \tag{20}$$

Simple calculations show that (20) can be derived from the following two relations:

$$\int_{a_m}^{\infty} \exp\{-C_1(m-1)x^{-\alpha}\} x^{-\beta-1} \, \mathrm{d}x \sim \frac{\Gamma(\beta/\alpha)}{\alpha C_1^{\beta/\alpha}} m^{-\beta/\alpha}, \tag{21}$$

$$\int_{a_m}^{\infty} \left\{ \tilde{F}^{m-1}(x) - \exp\{-C_1(m-1)x^{-\alpha}\} \right\} x^{-\beta-1} \, \mathrm{d}x = o\left(m^{-\beta/\alpha} \right). \tag{22}$$

Making a change of variables $y = C_1(m-1)x^{-\alpha}$ one has

$$\begin{aligned} &\int_{a_m}^{\infty} \exp\{-C_1(m-1)x^{-\alpha}\} x^{-\beta-1} \, \mathrm{d}x \\ &= \frac{1}{\alpha C_1^{\beta/\alpha}} (m-1)^{-\beta/\alpha} \left(\Gamma(\beta/\alpha) - \Gamma(\beta/\alpha, C_1(m-1)a_m^{-\alpha}) \right), \end{aligned}$$

where $\Gamma(\cdot)$ is a standard gamma function and

$$\Gamma(\alpha, x) = \int_x^\infty t^{\alpha-1} e^{-t}\, \mathrm{d}t$$

is an upper incomplete gamma function. Keeping in mind that $a_m^{-\alpha} m \to \infty$, we have

$$\Gamma(\beta/\alpha, C_1(m-1)a_m^{-\alpha}) \le \frac{1}{C_1(m-1)a_m^{-\alpha}} \Gamma\left(1 + \frac{\beta}{\alpha}\right) \to 0, \quad (m \to \infty).$$

This ends the proof of (21).

To prove (22), consider the difference $\Delta_m(x) := \tilde{F}^m(x) - \exp\{-C_1 m x^{-\alpha}\}$. We assume that m is large enough that inequalities

$$0 < C_1 a_m^{-\alpha} + C_2 a_m^{-\beta} < 1/2, \quad 1 + \frac{C_2\beta}{C_1\alpha} a_m^{-\beta+\alpha} > 0 \tag{23}$$

are satisfied. We recall that only C_1 is supposed positive, while C_2 may be negative. The second inequality in (23) ensures monotonous decay of a function $1 - \tilde{F}(x)$, $x \ge a_m$ and implies $0 < 1 - \tilde{F}(x) < 1/2$ for $x \ge a_m$. Then we can write :

$$\ln \tilde{F}(x) = -C_1 x^{-\alpha} - C_2 x^{-\beta} + r(x),$$

where $|r(x)| \le 2\left(C_1 + |C_2|\right)^2 x^{-2\alpha}$. Using this relation, for sufficiently large m, we have

$$\begin{aligned} |\Delta_m(x)| &= \exp\{-C_1 m x^{-\alpha}\} \left| \exp\left\{C_1 m x^{-\alpha} + m \ln \tilde{F}(x)\right\} - 1 \right| \\ &= \exp\{-C_1 m x^{-\alpha}\} \left| \exp\left\{-m\left(C_2 x^{-\beta} - r(x)\right)\right\} - 1 \right| \\ &\le C m x^{-(2\alpha\wedge\beta)} \exp\{-C_1 m x^{-\alpha}\}. \end{aligned}$$

Consequently, left hand side of (22) does not exceed

$$m \int_{a_m}^\infty \exp\{-C_1(m-1)x^{-\alpha}\} x^{-(2\alpha\wedge\beta)-\beta-1}\, \mathrm{d}x = o\left(m^{1-\frac{(2\alpha\wedge\beta)+\beta}{\alpha}}\right). \tag{24}$$

If $2\alpha \ge \beta$, then r.h.s. of (24) is $o\left(m^{1-2\beta/\alpha}\right) = o\left(m^{-\beta/\alpha}\right)$, while, in the case $2\alpha < \beta$, we have $o\left(m^{1-(2\alpha\wedge\beta)/\alpha-\beta/\alpha}\right) = o\left(m^{-\beta/\alpha-1}\right)$. Thus, (22) and, consequently, (20), are proved.

To finish the proof of (17), it remains to prove that the remainder term R_m from (19) is negligible, that is, $R_{m,i} = o\left(m^{-\zeta}\right)$, $i = 1, 2$. We have

$$|R_{m,1}| \leq Cm \int_{a_m}^{\infty} \left| F^{m-1}(x) - \tilde{F}^{m-1}(x) \right| x^{-\alpha-1} \, \mathrm{d}x.$$

The remainder term in (3) denote by $h(x)$, that is, $h(x) = 1 - F(x) - C_1 x^{-\alpha} - C_2 x^{-\beta} = x^{-\beta} h_1(x)$ where $h_1(x) = o(1)$, as $x \to \infty$. Let us rewrite the difference in the integrand as follows

$$F^{m-1}(x) - \tilde{F}^{m-1}(x) = \tilde{F}^{m-1}(x) \left(\exp\left\{ (m-1) \ln\left(1 - \frac{h(x)}{\tilde{F}(x)}\right) \right\} - 1 \right).$$

One can assume that for $x \geq a_m$ inequality $\left| h(x)/\tilde{F}(x) \right| < 1/2$ is satisfied, thus from the Taylor expansion of $\ln\left(1 - h(x)/\tilde{F}(x)\right)$ it follows that there exist a constant $C > 0$ such that $|\ln\left(1 - h(x)/\tilde{F}(x)\right)| \leq C|h(x)|/\tilde{F}(x)$. Since for $x \geq a_m$ the product $(m-1)x^{-\beta}$ tends to zero as $m \to \infty$, we can assume that inequality

$$(m-1) \left| \ln\left(1 + \frac{h(x)}{\tilde{F}(x)}\right) \right| < 1$$

holds. Then, by applying inequality $|\mathrm{e}^x - 1| \leq C|x|$, $0 < |x| < 1$, we get

$$\begin{aligned} \left| F^{m-1}(x) - \tilde{F}^{m-1}(x) \right| &\leq (m-1)\tilde{F}^{m-1}(x) \left| \ln\left(1 - \frac{h(x)}{\tilde{F}(x)}\right) \right| \\ &\leq C(m-1)\tilde{F}^{m-1}(x) \, |h(x)| \, . \end{aligned}$$

Applying the last inequality we have

$$\begin{aligned} R_{m,1} &\leq Cm \int_{a_m}^{\infty} \tilde{F}^{m-1}(x) \, |h(x)| \, x^{-\alpha-1} \, \mathrm{d}x \\ &\leq Cm a_m^{-\alpha} \sup_{x \geq a_m} |h_1(x)| \int_{a_m}^{\infty} x^{-\beta-1} \tilde{F}^{m-1}(x) \, \mathrm{d}x. \end{aligned}$$

Taking into account (21)-(22) we obtain $R_{m,1} \leq C a_m^{-\alpha} m^{-\zeta} \sup_{x \geq a_m} |h_1(x)| = o\left(m^{-\zeta}\right)$. In a similar way we get

$$|R_{m,2}| \leq Cm^{-\zeta} \sup_{x \geq a_m} |h(x)| = o\left(m^{-\zeta}\right)$$

and the proof of (17) is completed.

To complete the proof of (8) it remains to prove (14). The random variables $\kappa_{n,1}, \dots, \kappa_{n,n}$, defined in (2), are i.i.d.. Therefore $\mathbf{E}\,\hat{p} = \mathbf{E}\,\kappa_{n,1}$. Now it is clear that

$$
\begin{aligned}
\mathbf{E}\,\hat{p} &= m! \int_0^\infty \frac{\mathrm{d}F(x_1)}{x_1} \int_0^{x_1} x_2 \,\mathrm{d}F(x_2) \int_0^{x_2} \mathrm{d}F(x_2) \dots \int_0^{x_{m-1}} \mathrm{d}F(x_m) \\
&= (m-1)m \int_0^\infty \left\{ \int_0^{x_1} x_2 F^{m-2}(x_2)\,\mathrm{d}F(x_2) \right\} \frac{\mathrm{d}F(x_1)}{x_1}.
\end{aligned}
$$

Integrating by parts the inner integral we get

$$
\begin{aligned}
\mathbf{E}\,\hat{p} &= m \int_0^\infty \left\{ x_1 F^{m-1}(x_1) - \int_0^{x_1} F^{m-1}(x_2)\,\mathrm{d}x_2 \right\} \frac{\mathrm{d}F(x_1)}{x_1} \\
&= 1 - m \int_0^\infty \left\{ \int_0^{x_1} F^{m-1}(x_2)\,\mathrm{d}x_2 \right\} \frac{\mathrm{d}F(x_1)}{x_1}.
\end{aligned}
$$

It remains to change order of integration to conclude the proof of (14).

Proof of (11). Since the proofs of (11) and (7) are essentially the same, we shall give main steps only. In view of decomposition

$$
\sqrt{n}\,(\hat{p} - p) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (\kappa_{n,i} - \mathbf{E}\,\kappa_{n,i}) + \sqrt{n}\gamma_m,
$$

the assertion (11) follows from

$$
\frac{1}{\sqrt{n}} \sum_{i=1}^n (\kappa_{n,i} - \mathbf{E}\,\kappa_{n,i}) \xrightarrow{\mathrm{D}}_{N\to\infty} N(0, \sigma^2), \tag{25}
$$

$$
\sqrt{n}\gamma_m \to \mu, \quad (m = m_{opt},\ N \to \infty). \tag{26}
$$

To prove (25) check Lyapunov condition for i.i.d. random variables forming triangular array, while relation (26) one can verify by using (8) and (9) with $m = m_{opt}$.

Proof of (10). From (11) we know that the variance $\mathbf{E}\,(\hat{p} - \mathbf{E}\,\hat{p})^2$ asymptotically equals $\sigma^2 m/N$. Taking the main term in the asymptotic relation (8) we get that the asymptotic mean squared error of $\hat{p}$ equals $\chi^2 m^{-2\zeta} + \sigma^2 m/N$.

Since this is a very simple function with respect to m it is easy to see that a solution of a minimization problem

$$\inf_{2\le m\le N}\left\{\chi^2 m^{-2\zeta}+\sigma^2 m/N\right\} \tag{27}$$

is given by (9). Here it is necessary to note that we require that N is sufficiently large, since for values of ζ close to 0 the solution of minimization problem (for a given N) may be smaller than 2. Also one can note that instead of taking main term from (8) we can take the sequence γ_m and apply Lemma 2.8 in Dekkers and de Haan (1993), this will give the same result. Having (9)one can easily get (10). The theorem is proved.

3 Comparison of estimators

In this section we compare the tail index estimator $\hat{p}$ with the estimators $\gamma_{N,k}^{(j)}$, $j=1,2,3,4$, using the same method as in de Haan and Peng (1998). May be here it is worth to mention that in a recent paper Nematollahi and Tafakori (2007) there was proposed another approach to compare tail index estimators, but this method of comparison is well adapted to a specific estimator introduced by fan in fan (2004). We recall that $\hat{p}$ estimates the quantity $p=\alpha/(\alpha+1)$, while the four above mentioned estimators estimate $\gamma=1/\alpha$, that is, different function of the unknown parameter α. Therefore the first step in comparison is to transfer the estimators to the same function, and we had chosen to compare the estimator $\hat{p}$ with estimators $p_{N,k}^{(j)}=1/(1+\gamma_{N,k}^{(j)})$. We need the following simple statement.

Lemma 4. *Suppose (3) holds. Let $k^{(j)}=k^{(j)}(N)$, $j=1,2,3,4$ be a sequences of integers with*

$$k^{(j)}(N)\to\infty \quad \text{and} \quad k^{(j)}(N)/N\to 0, \quad (N\to\infty) \tag{28}$$

and let estimators $\gamma_{N,k}^{(j)}$ are asymptotically normal, i.e. there exist constants $b_j\in\mathbb{R}$ and $\sigma_j>0$ such that

$$\sqrt{k^{(j)}}\left(\gamma_{N,k}^{(j)}-\gamma\right)\xrightarrow{\mathrm{D}}_{N\to\infty}\mathcal{N}(b_j,\sigma_j^2). \tag{29}$$

Then

$$\sqrt{k^{(j)}}\left(p_{N,k}^{(j)}-p\right)\xrightarrow{\mathrm{D}}_{N\to\infty}\mathcal{N}\left(-\frac{b_j}{(1+\gamma)^2},\frac{\sigma_j^2}{(1+\gamma)^4}\right). \tag{30}$$

Proof. We use the obvious identity

$$\sqrt{k^{(j)}}\left(p_{N,k}^{(j)}-p\right)=\frac{\sqrt{k^{(j)}}\left(\gamma-\gamma_{N,k}^{(j)}\right)}{(1+\gamma)(1+\gamma_{N,k}^{(j)})}.$$

Now relation (30) follows from this identity, Theorem 4.4 in Billingsley (1968), relation (29) and the relation $\gamma_{N,k}^{(j)} \xrightarrow{\mathrm{P}} \gamma$, $(N\to\infty)$. The lemma is proved.

De Haan and Peng proved (see Thm.2 in de Haan and Peng (1998)) that the asymptotic second moment of $\gamma_{N,k}^{(j)}-\gamma$ is minimal and equals $\left(k^{(j)}\right)^{-1}\sigma_k^2(1+2\zeta)/(2\zeta)$, if $k^{(j)}$ satisfies the relation

$$\lim_{N\to\infty} k^{(j)}A^2(N/k^{(j)})=\frac{\sigma_j^2}{2\zeta D_k^2},\quad j=1,2,3,4, \tag{31}$$

where

$$\begin{aligned} D_1 &= \frac{1}{1+\zeta},\ D_3=\frac{1}{1+\zeta}-\frac{\alpha\zeta}{(1+\zeta)^2},\\ D_2 &= \frac{1}{(2^{1/\alpha}-1)\ln 2}\frac{1-2^{-\zeta}}{\zeta}\left(2^{(1/\alpha)-\zeta}-1\right),\ D_4=\frac{1}{(1+\zeta)^2}, \end{aligned}$$

and

$$\sigma_1^2=\frac{1}{\alpha^2},\ \sigma_2^2=\frac{1+2^{2/\alpha+1}}{\alpha^2(2^{1/\alpha}-1)^2\ln^2 2},\ \sigma_3^2=\frac{1+\alpha^2}{\alpha^2},\ \sigma_4^2=\frac{2}{\alpha^2}$$

are limit variances in (29). The function $A(t)$ in (31) has the asymptotic

$$A(t)\sim-\frac{\zeta}{\alpha}\frac{C_2}{C_1^{\beta/\alpha}}t^{-\zeta},\quad t\to\infty.$$

We recall that the function $A(t)$ was introduced in (4).

Denote by $k_{opt}^{(j)}$ a sequence $k^{(j)}$, satisfying (31). From (31), we have

$$\begin{aligned}
k_{opt}^{(1)}(N) &\sim \left(\frac{(1+\zeta)^2}{2\zeta^3} \frac{(C_1)^{2\beta/\alpha}}{(C_2)^2} \right)^{1/(1+2\zeta)} N^{2\zeta/(1+2\zeta)}, \\
k_{opt}^{(2)}(N) &\sim \left(\frac{1+2^{2/\alpha+1}}{2\zeta(1-2^{-\zeta})^2(2^{1/\alpha-\zeta}-1)^2} \frac{(C_1)^{2\beta/\alpha}}{(C_2)^2} \right)^{1/(1+2\zeta)} N^{2\zeta/(1+2\zeta)}, \\
k_{opt}^{(3)}(N) &\sim \left(\frac{(1+\zeta)^4(1+\alpha^2)}{2\zeta^3(1+\zeta-\zeta\alpha)^2} \frac{(C_1)^{2\beta/\alpha}}{(C_2)^2} \right)^{1/(1+2\zeta)} N^{2\zeta/(1+2\zeta)}, \\
k_{opt}^{(4)}(N) &\sim \left(\frac{(1+\zeta)^4}{\zeta^3} \frac{(C_1)^{2\beta/\alpha}}{(C_2)^2} \right)^{1/(1+2\zeta)} N^{2\zeta/(1+2\zeta)}.
\end{aligned}$$

Under normalization $k_{opt}^{(j)}(N)$ instead of $k^{(j)}(N)$ in (29) the limit random variable has a mean

$$\frac{\sigma_k}{\sqrt{2\zeta}} \operatorname{sgn} \left(D_k \lim_{N\to\infty} \sqrt{k_{opt}^{(j)}(N)} A\left(N/k_{opt}^{(4)}(N)\right)\right).$$

Moreover, Lemma 4 imply

$$\mathbf{E}\left(p_{N,k}^{(j)} - p\right)^2 \sim \left(\frac{\alpha}{\alpha+1}\right)^4 \frac{2\beta-\alpha}{2(\beta-\alpha)} \frac{\sigma_j^2}{k_{opt}^{(j)}(N)}, \quad (N\to\infty). \tag{32}$$

Now it is possible to compare the estimator $\hat{p}$ with the estimators $p_{N,k}^{(j)}$ as in was done in de Haan and Peng (1998), i.e., by calculating a limit of the ratio of minimal mean squared errors:

$$RMMSE(j) = \lim_{N\to\infty} \frac{\mathbf{E}\,(\hat{p}-p)^2}{\mathbf{E}\left(p_{N,k}^{(j)}-p\right)^2}.$$

From (10) and (32) we have the following results:

$$\begin{aligned}
RMMSE(1) &= \left(\eta(\alpha,\beta)\Gamma^2(2+\zeta)\right)^{1/(1+2\zeta)}, \\
RMMSE(2) &= (2^{(1/\alpha)}-1)^2 \ln^2(2)\Bigg(\eta(\alpha,\beta)\times \\
&\quad \times\frac{\zeta^2\Gamma^2(1+\zeta)}{(1-2^{-\zeta})^2(2^{(2/\alpha)+1}+1)^{2\zeta}(2^{1/\alpha-\zeta}-1)^2}\Bigg)^{1/(1+2\zeta)}, \\
RMMSE(3) &= \left(\eta(\alpha,\beta)\frac{(1+\zeta)^2\Gamma^2(2+\zeta)}{(1+\alpha^2)^{2\zeta}(1+\zeta-\alpha\zeta)^2}\right)^{1/(1+2\zeta)}, \\
RMMSE(4) &= \left(\eta(\alpha,\beta)\frac{(1+\zeta)^2\Gamma^2(2+\zeta)}{2^{2\zeta}}\right)^{1/(1+2\zeta)},
\end{aligned}$$

where

$$\eta(\alpha,\beta)=\left(\frac{\beta(\alpha+1)}{\alpha(\beta+1)}\right)^2\left(\frac{(\alpha+1)^2}{\alpha(\alpha+2)}\right)^{2\zeta}.$$

It is easy to conclude that $RMMSE(1) > 1$ for all $0 < \alpha < \beta$ i.e., Hill estimator $p_{N,k}^{(1)}$ dominates estimator $\hat{p}$. Due to the inequality $(\alpha+1)^6 - 4\alpha^3(\alpha+2) > 0$, for $\alpha > 0$ (it follows from the binomial formula), the same conclusion is valid for de Vries estimator $p_{N,k}^{(4)}$.

Comparison of estimators $\hat{p}$, $p_{N,k}^{(2)}$ and $p_{N,k}^{(3)}$ is shown in Figures 1a-1c. α values are on the horizontal axis, while vertical axis labels β values. In all three figures the area $\{(\alpha,\beta):\ 0<\beta<\alpha\}$ (those values of parameters that are not considered) is left as white. In Figure 1a the area $\{(\alpha,\beta):\ RMMSE(2)>1\}$ is in black and $\{(\alpha,\beta):\ RMMSE(2)<1\}$ is in dark grey. Similarly, Figure 1b presents comparison of the estimators $\hat{p}$ and $p_{N,k}^{(3)}$: as in Figure 1a, the area $\{(\alpha,\beta):\ RMMSE(3)<1\}$ is in dark grey and the area $\{(\alpha,\beta):\ RMMSE(3)>1\}$ - in light grey. Finally, Figure 1c gives areas of domination estimators $\hat{p}$ (dark grey), $p_{N,k}^{(2)}$ (black), and $p_{N,k}^{(3)}$ (light grey).

As it was mentioned in the Introduction, in de Haan and Peng (1998) the comparison of the estimators $\gamma_{N,k}^{(j)}$, $j=1,2,3,4$ was performed with respect to the parameters (γ,ρ). For the sake of completeness we include analogous of the Figures 1a-1c in the plane (ρ,γ) also. In the Figures 2a-2c the horizontal axis labels γ values, while vertical axis labels ρ values. As in Figures 1a-1c, the area where estimator $\hat{p}$ has an asymptotic mean

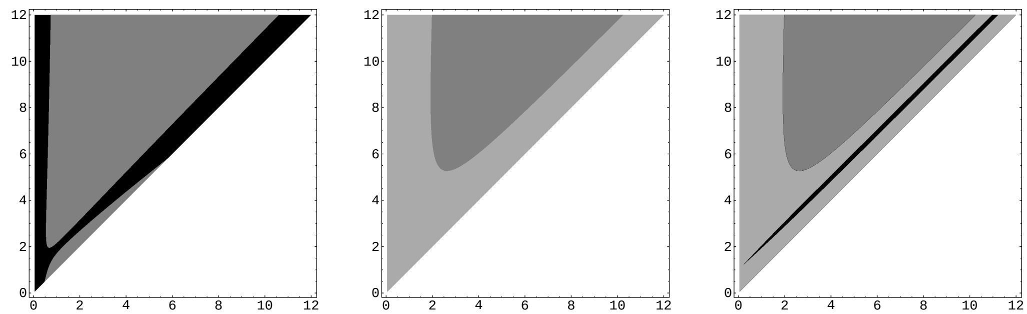

Figure 1: 1a, 1b and 1c figures

squared error smaller than the other estimator(s) is in dark grey. A black and light grey colors mark the areas of domination of estimators $p^{(2)}_{N,k}$ or $p^{(3)}_{N,k}$, respectively.

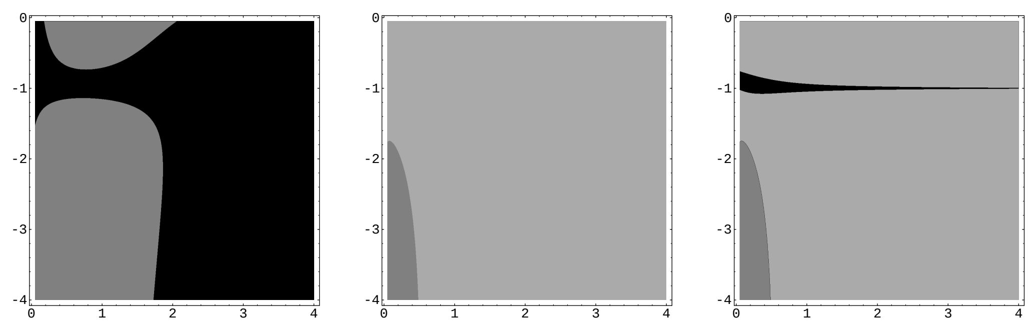

Figure 2: 2a, 2b and 2c figures

References

Billingsley, P. (1968), *Convergence of Probability Measures*, Wiley, New York.

Davydov, Yu. and V. Paulauskas (1999), On the estimation of the parameters of multivariate stable distributions, *Acta Appl. Math.*, 58, 107–124.

Davydov, Yu., V. Paulauskas and A. Račkauskas (2000), More on P-stable convex sets in Banach spaces, *J. Theoret. Probab.*, 13, 39–64.

De Haan, L. and L. Peng (1998), Comparison of tail index estimators, *Statist. Neerlandica*, 52, 60–70.

Dekkers, A.L.M. and L. de Haan (1993), Optimal choice of sample fraction in extreme-value estimation, *J. Multivariate Anal.*, 47, 173–195.

Dekkers, A.L.M., J.H.J. Einmahl and L. de Haan (1989), A moment estimator for the index of an extreme-value distribution, *Ann. Statist.*, 17, 1833-1855.

Fan, Zh. (2004), Estimation problems for distributions with heavy tails, *J. Statist. Plann. Inference*, 123, 13–40.

Fraga Alves, M.I. (2001), A location invariant Hill-type estimator, *Extremes*, 4, 199–217.

Fraga Alves, M.I., L. de Haan and T. Lin (2006), Third order extended regular variation, *Publications de lInstitut Mathematique*, 80, 109–120.

Fraga Alves, M.I., M.I. Gomes and L. de Haan (2003), A new class of semiparametric estimators of the second order parameter, *Port. Math.*, 60, 194–213.

Gadeikis, K. and V. Paulauskas (2005), On the estimation of a change point in a tail index, *Lith. Math. J.*, 45, 272–283.

Geluk, J.L. and L. Peng (2000), An adaptive optimal estimate of the tail index for MA(1) time series, *Statist. Probab. Lett.*, 46, 217–227.

Gomes, M.I., L. de Haan and L. Henriques (2008), Tail index estimation through the accommodation of bias in the weighted log-excesses, *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 70, 31–53.

Gomes , M.I. and M.J. Martins (2002), Asymptotically unbiased estimators of the tail index based on external estimation of the second order parameter, *Extremes*, 5, 5–31.

Hall, P. (1982), On some simple estimates of an exponent of regular variation, *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 44, 37–42.

Hill, B.M. (1975), A simple general approach to inference about the tail of a distribution, *Ann. Statist.*, 3, 1163–1174.

LePage, R., M. Woodroofe and J. Zinn (1981). Convergence to a stable distribution via order statistics. *Ann. Prob.* 9, 624–632.

Li, J., Z. Peng and S. Nadarajah (2008), A class of unbiased location invariant Hill-type estimators for heavy tailed distributions, *Electron. J. Stat.*, 2, 829–847.

Markovich, N. (2007), *Nonparametric Analysis of Univariate Heavy-Tailed Data*, Jon Wiley & Sons, Chichester.

Nematollahi, A. R. and L. Tafakori (2007), On Comparison of the Tail Index of Heavy Tailed Distributions Using Pitman's Measure of Closeness, *Appl. Math. Sci.*, 1, 909–914.

Paulauskas, V. (2003), A New Estimator for a Tail Index, *Acta Appl. Math.* , 79, 55–67.

Paulauskas, V. and M. Vaičiulis (2010), Some new modifications of DPR estimator of the tail index, *preprint*.

Pickands, J. (1975), Statistical inference using extreme order statistics, *Ann. Statist.* , 3, 119–131.

Qi, Y. (2010), On the tail index of a heavy tailed distribution, *Ann. Inst. Statist. Math.*, 62(2), 277–298.

Quintos, C., Zh. Fan and P. Phillips (2001),Structural change tests in tail behavior and the Asian crisis, *Rev. Econom. Stud.*, 13, 633–663.

Resnick, S. and C. Starica (1997), Smoothing the Hill estimator, *Adv. in Appl. Probab.*, 29, 271–293.

Smith, R.L. (1987), Estimating tails of probability distributions, *Ann. Statist.*, 15, 1174–1207.

Weissman, I. (1978), Estimation of parameters and large quantiles based on the k largest observations, *Journal of American Statistical Association* , 73, 812–815.